2-2014

Cursing in English on Twitter

Wenbo Wang
Wright State University - Main Campus

Lu Chen
Wright State University - Main Campus

Krishnaprasad Thirunarayan
Wright State University - Main Campus, t.k.prasad@wright.edu

Amit P. Sheth
Wright State University - Main Campus, amit@sc.edu

Repository Citation

Wang, W., Chen, L., Thirunarayan, K., & Sheth, A. P. (2014). Cursing in English on Twitter. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 415-424.
https://corescholar.libraries.wright.edu/knoesis/590

Cursing in English on Twitter

Wenbo Wang, Lu Chen, Krishnaprasad Thirunarayan, Amit P. Sheth
Kno.e.sis Center, Wright State University
Dayton, OH 45435 USA
{wenbo, chen, tkprasad, amit}@knoesis.org

ABSTRACT
Cursing is not uncommon during conversations in the physical world: 0.5% to 0.7% of all the words we speak are curse words, given that 1% of all the words are first-person plural pronouns (e.g., we, us, our). On social media, people can instantly chat with friends without face-to-face interaction, usually in a more public fashion and broadly disseminated through highly connected social network. Will these distinctive features of social media lead to a change in people's cursing behavior? In this paper, we examine the characteristics of cursing activity on a popular social media platform – Twitter, involving the analysis of about 51 million tweets and about 14 million users. In particular, we explore a set of questions that have been recognized as crucial for understanding cursing in offline communications by prior studies, including the ubiquity, utility, and contextual dependencies of cursing.

Author Keywords
Profanity; Social Media; Twitter; Emotion; Gender Difference; Cursing.

ACM Classification Keywords
J.4 SOCIAL AND BEHAVIORAL SCIENCES: Psychology

General Terms
Human Factors; Experimentation; Verification.

INTRODUCTION
Do you curse? Do you curse on social media? How often do you see people cursing on social media (e.g., Twitter)? Cursing, also called swearing, profanity, or bad language, is the use of certain words and phrases that are considered by some to be rude, impolite, offensive, obscene, or insulting [26]. In this paper, we use cursing, profanity and swearing interchangeably. As Jay [6] pointed out, cursing is a "rich emotional, psychological and sociocultural phenomenon", which has attracted many researchers from related fields such as psychology, sociology, and linguistics [7, 8].

CSCW'14 , February 15 - 19 2014, Baltimore, MD, USA

ACM 978-1-4503-2540-0/14/02$15.00.
http://dx.doi.org/10.1145/2531602.2531734

Over the last decade, social media has become an integral part of our daily lives. According to the 2012 Pew Internet & American Life Project report [25], 69% of online adults use social media sites and the number is steadily increasing. Another Pew study in 2011 [24] shows that 95% of all teens with ages 12-17 are now online and 80% of those online teens are users of social media sites. People post on these sites to share their daily activities, happenings, thoughts and feelings with their contacts, and keep up with close social ties, which makes social media both a valuable data source and a great target for various areas of research and practice, including the study of cursing. While the CSCW community has made great efforts to study various aspects (e.g., credibility [13], privacy [1]) of social networking and social media, our understanding of cursing on social media still remains very limited.

The communication on social media has its own characteristics which differentiates it from offline interaction in the physical world. Let us take Twitter for example. The messages posted on Twitter (i.e., tweets) are usually public and can spread rapidly and widely through the highly connected user network, while the offline conversations usually remain private among the persons involved. In addition, we may find that more of our actual exchange of words in the physical world happens through face-to-face oral communication, while on Twitter we mostly communicate by writing/typing without seeing each other. Will such differences lead to a change in people's cursing behavior? Will the existing theories on swearing during the offline communication in physical world still be supported if tested on social media?

To address such differences, this paper examines the use of English curse words on the micro-blogging platform Twitter. We collected a random sampling of all public tweets and the data of relevant user accounts every day for four weeks. We first identified English cursing tweets in the collection, and extracted numerous attributes that characterize users and users' tweeting behaviors. We then evaluated the effect of these attributes with respect to the cursing behaviors on Twitter. This exploratory study aims to improve our understanding of cursing on social media by exploring a set of questions that have been identified as crucial in previous cursing research on offline communication. The answers to these questions may also have valuable implications for the studies of language acquisition, emotion, mental health, verbal abuse, harassment, and gender difference [6].

Specifically, we examine four research questions:

- Q1 (*Ubiquity*): How often do people use curse words on Twitter? What are the most frequently used curse words?
- Q2 (*Utility*): Why do people use curse words on Twitter? Previous studies [6] found that the main purpose of cursing is to express emotions. Do people curse to express emotions on Twitter? What are the emotions that people express using curse words?
- Q3 (*Contextual Variables*): Does the use of curse words depend on various contextual variables such as time (when to curse), location (where to curse), or communication type (how to curse)?
- Q4: Who says curse words to whom on Twitter? Previous research [5, 11] suggested that gender and social rank of people play important roles in cursing; do they also affect people using or hearing curse words on Twitter?

The remainder of the paper proceeds as follows. We first review the related work in the next section. We then describe the data and methods, including an overview of the data and how we obtained it, the methods to identify cursing tweets and to extract various attributes, the exploration of the above mentioned questions, and the implications of the results. Finally, we close with a summary of this study.

RELATED WORK

We first consider the previous research on cursing in offline communications, and organize them into four groups corresponding to the four questions we address in this study. Then we discuss the efforts that have been made to understand the cursing phenomenon on social media.

The Ubiquity of Cursing

Cursing is more common than people might think. Jay [4, 6] found 70 curse words in an 11,609-word tape-recorded conversation of elementary school students and college students. In another study [12], Mehl and Pennebaker reported that curse words occurred at a rate of 0.5% over two 2-day periods among undergraduate students, which may not seem significant except that the first person plural pronouns – words like *we*, *us*, and *our* – occurred at a 1.0% rate. They also found substantial differences among individuals regarding curse words usage: the word rates varied from a minimum of 0% to a maximum of 3.4%. Some recent studies suggest that people have been hearing and using profanity more often than ever before [27], and more (a 69% increase) and harsher curse words have been used in TV programs in 2010 compared to 2005 [22]. Researchers also found that a few most frequently used words (e.g., *fuck*, *damn*, *hell*, and *shit*) account for most of the cursing expressions in conversations (a long tail phenomenon)[4].

The Utility of Cursing

Cursing is not as negative or harmful as it may seem at first glance. Prior studies [4, 5, 11, 14] suggest that the main reason that people use swearing words is to express some strong emotions, especially anger and frustration, for emphasis. As a common conversational practice, cursing rarely results in obvious harm. Only when cursing occurs in the form of insults toward others, such as name calling, harassment, hate speech, and obscene telephone calls, it becomes harmful [7]. Researchers also found other positive effects could be achieved by swearing. For example, Stephens et al. [17] reported that swearing increased pain tolerance and decreased perceived pain compared with not swearing. In addition, people may find relief and positive effects of laughing at jokes, humour and sarcasm in which curse words are used [5, 7].

Contextual Variables

Prior studies suggest that cursing is very sensitive to contextual influences [5]. More specifically, people's propensity to curse, the particular curse words people use, and how others perceive the cursing behavior, are dependent on various contextual variables. Generally, the context of cursing activity is defined by those variables about *when*, *where*, *how*, *who* and *with whom* the cursing occurs. Among those variables, while *who* and *with whom* variables have attracted the most attention from researchers, physical location has also been recognized as important [4, 5, 8]. Researchers [8] found that "people are more likely to swear in relaxed environments than in formal environments" (e.g., pub vs. office). Since such observations were made in the setting of oral communication in physical world, it is not clear whether physical location still matters for cursing on social media which occurs as written messages in digital world. In addition, little is known about how *when* and *how* factors would affect cursing. In this study, we examine the effect of location variable as well as the variables of time during a day, day in a week and message types, on cursing on social media.

Who Says Curse Words to Whom

There have been a considerable number of studies on understanding the characteristics of people who use and hear curse words. A set of important variables have been identified and investigated [4, 5, 11], including gender, age, race, religion and power. Unfortunately, many of these variables such as age, race and religion remain difficult to measure on Twitter, thus, we limit our focus on gender and power. Some well-recognized patterns about gender in swearing research include: (1) Gender affects cursing frequency. Many studies [9, 12, 19, 15] suggest that men curse more frequently than women. (2) Gender affects the choice of curse words. For example, according to a study in 2006 [11], women used the words *god*, *hell*, *bitch*, and *piss* more often than men, and men used the words *fuck* and *cunt* more often than women. (3) People are more likely to curse in same-gender contexts than in mixed-gender contexts [8, 15]. People's power or social rank also plays a role in cursing. McEnery [11] found that frequency of cursing is inversely proportional to the social rank.

Cursing on social Media

Only a few efforts have been made to explore cursing on social media. Thelwall [19] studied the use of curse words in MySpace profiles and the effects of gender and age factors. Sood et al. [16] investigated profanity usage in Yahoo! Buzz communities and found that different communities (e.g., politics or sports) use profanity with varied frequencies and in different ways. Turning to Twitter, Bak et al. [2] studied self-disclosure behavior in Twitter conversations. As one aspect of self-disclosure, profanity was used more frequently between users with higher relationship strength. While other researchers have mainly focused on investigating algorithms to automatically detect offensive tweets [30, 31], our understanding of the basic questions regarding the use of offensive words on Twitter still remains unexplored, such as why people use curse words, who uses it, and whether these words are always harmful and should be removed. The insights gained in this study can shed light on these questions.

METHOD AND ANALYSIS

Data Collection

Twitter provides a small random sample of all public tweets via its *sample API* in real time [1]. Using this API, we continuously collected tweets for four weeks from March 11th 2013 to April 7th 2013. We kept only the users who specified 'en' as their language in profiles. Further, we utilized Google Chrome Browser's embedded language detection library to remove non-English tweets [2]. In total, we gathered about 51M tweets from 14M distinct user accounts.

Spam on Twitter may impede the delivery of quality results from data analysis. To examine the spammers in our dataset, a random set of 200 user accounts were selected and manually verified based on the content of tweets and their profile (using the number of friends, followers, etc.) attached with each account. Of the 200 accounts, 5 (2.5%) were identified as spammers, and there were 88 tweets in our dataset from these 5 spammers, accounting for 1.32% of all 6678 tweets posted by these 200 users. On the other hand, we observed that there were some accounts that posted suspiciously frequently, and it could harm our analysis if they were spammers. Thus, we manually verified the top 1000 accounts which posted most frequently in our dataset, and removed the identified spam accounts and their tweets. Not surprisingly, among the 1000 accounts, there were 19 spammers in the top 100 accounts, 15 spammers in the following 100 accounts, and then this fraction kept diminishing, with only 3 spammers identified in each of the last two sets of 100 accounts. Totally, we removed 68 spammers and 89,556 tweets from our dataset.

Cursing Lexicon Coding

To create a lexicon of curse words for this study, we first collected existing curse word lists from Internet used by native English speakers for cursing on social media. Besides the curse word list [23] that have been used by existing studies [16, 30], we collected additional curse word lists from [28, 21, 20] to increase the coverage. After merging the above word lists, we found that a few non-curse words were also included, e.g., *sexy*. Also, there are some non-English words, e.g., *buceta*, which means *pussy* in Portuguese. Moreover, some words can be used in both cursing and non-cursing contexts: *gay* in *"you are so gay"* conveys cursing, but gay in *"Bill Clinton urges Illinois to approve gay marriage bill"* does not convey cursing . To achieve high precision in identifying cursing expressions, we keep only the words that are mostly associated with cursing connotation.

[1] `https://dev.twitter.com/docs/api/1.1/get/statuses/sample`

[2] `https://pypi.python.org/pypi/chromium_compact_language_detector/0.2`

Statistics	Min	Max	Mean	Median	Std. Dev.
Overall Tweets	1.0	4124.0	3.56	2.0	8.00
Cursing Tweets	1.0	549.0	1.78	1.0	2.39

Table 1. Statistics of overall tweets and cursing tweets per user

Specifically, to retain these curse words, we asked two college students who are native English speakers to independently annotate the collected words in the context of social media with the following labels: 1 - the word is mostly used for cursing, 2 - the word can be used for both cursing and non-cursing purposes, 3 - usually the word is not used for cursing, or 4 - I do not know its meaning. Cohen's Kappa between the labels chosen by the two students was 0.5582. In the end, we kept only 788 words that received label '1' from both students to emphasize high precision. Besides correctly spelled words, (e.g., *fuck*, *ass*), the lexicon also included different variations of curse words, e.g., *a55*, *@$$*, *$h1t*, *b!tch*, *bi+ch*, *c0ck*, *f*ck*, *l3itch*, *p*ssy*, and *dik*.

We call a tweet *cursing tweet* if it contains at least one curse word. Twitter users may use different variations of the same word, so we first simply compare words in a tweet against all the curse words in the lexicon. If there is no match, we remove repeating letters in the words (e.g., fuckk → fuck) of a tweet and repeat the matching process. We also convert digits or symbols in a word to their original letters: e.g., 0 → o, 9 → g, ! → i. Moreover, based on our observations, the following symbols, '_', '%', '-', '.', '#', '\', ''', are frequently used to mask curse words: *f_ck*, *f%ck*, *f.ck*, *f#ck*, *f'ck* → *fuck*. We apply the edit distance approach similar to [16] to spot curse words with mask symbols. Namely, if the edit distance between a candidate word (*f_ck*) and a curse word (*fuck*) equals the number of mask symbols (1 in this case) in the candidate word, then it is a match. Table 1 provides an overview of the per-user count of the number of overall tweets and cursing tweets in our data collection.

To evaluate the accuracy of this lexicon-based method to spot cursing tweets, we drew a random sample of 1000 tweets, and asked two annotators to manually label them as cursing or non-cursing independently. Finally, there were 118 tweets labeled as cursing tweets for which both annotators agreed on their labels, and the other 882 tweets were labeled as non-cursing ones. We then tested the lexicon-based spotting approach on this labeled dataset, and the results showed that this lexicon-based method achieved a precision of 98.84%, a recall of 72.03% and F1 score of 83.33%. As expected, this lexicon-based approach for profanity detection provides high precision but lower recall, which is mainly due to the variations in curse words (e.g., due to misspellings and abbreviations) and context sensitivity of cursing. Though we believe that, for this work, high-precision is preferred and recall of 72.03% is considered reasonable, more sophisticated classification methods that can further improve the recall remain an interesting topic for future work.

Cursing Frequency and Choice of Curse Words

Prior studies have found that 0.5% to 0.7% of all the words we speak in our daily lives are curse words [4, 12]. Turning to Internet chatrooms, Subrahmanyam et. al. [18] reported that 3% of utterances contain curse words. Our comparison of cursing frequencies from different studies is shown in Table 2. Compared with existing studies, our estimate of cursing frequency was conducted for a significantly larger population: 14 million Twitter users and 51 million tweets. After removing punctuation marks and emoticons, we find that curse words occurred at the rate of 1.15% on Twitter, which is more than twice the rate (0.5%) in [12]. About 7.73% of all the tweets in our collection contain curse words, namely, one out of 13 tweets contains curse words. If we consider one tweet as roughly one utterance, this rate is also more than twice the rate (3%) in [18].

Besides the cursing frequency, we are also interested in the question: Which curse words are most popular? We manually grouped different variations of curse words into their root forms, e.g., *@$$*, *a$$*, $\rightarrow$ *ass*. If a curse word is the combination of two or more words, and one of its component words is also a curse word, then it will be grouped into its cursing component word, e.g., *dumbass*, *dumbasses*, *@sshole*, *a$$h0!e*, *a55hole* $\rightarrow$ *ass*. All the 788 curse words are grouped into 89 distinct groups based on the root curse words and the frequencies of the top 20 words are shown in Figure 1. The most popular curse word is *fuck*, which covers 34.73% of all the curse word occurrences, followed by *shit* (15.04%), *ass* (14.48%), *bitch* (10.34%), *nigga* (9.68%), *hell* (4.46%), *whore* (1.82%), *dick* (1.67%), *piss* (1.53%), and *pussy* (1.16%).

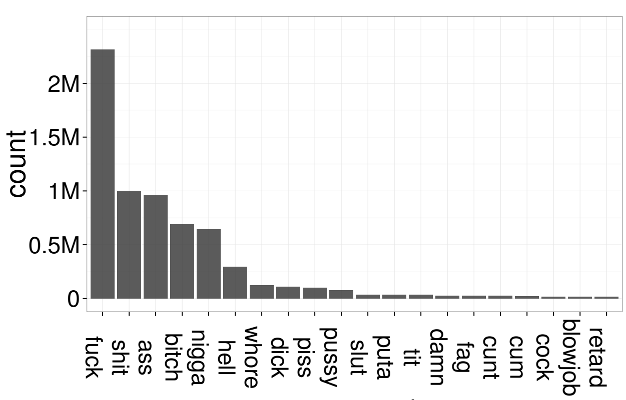

Figure 1. Frequencies of curse words: Only top 20 curse words are shown due to space limitation.

Realizing that only a small subset of curse words occurs very frequently, we also draw the cumulative distribution of top 20 curse words. We find that the top seven curse words – *fuck*, *shit*, *ass*, *bitch*, *nigga*, *hell* and *whore* cover 90.55% of all the curse word occurrences.

Cursing vs. Emotion

Psychology studies [8] suggest that "the main purpose of cursing is to express emotions, especially anger and frustration." Thus, we aim to explore emotions expressed in cursing

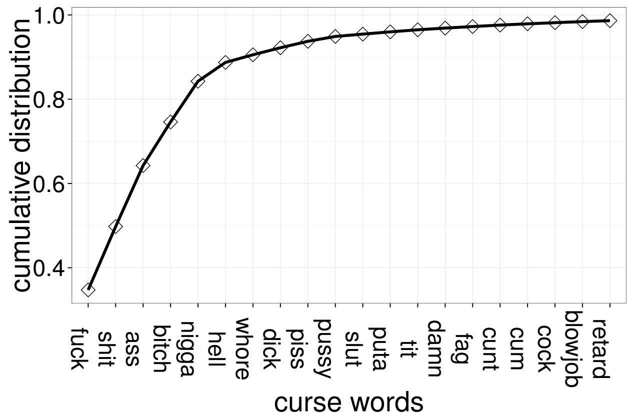

Figure 2. Cumulative distribution of curse Words: The top 7 curse words cover more than 90% of all the curse word occurrences.

tweets and compare them with those in non-cursing tweets. We adapted the emotion identification approach from our prior work [29] to automatically detect emotions expressed in tweets. The basic idea is to leverage ending emotion hashtags to automatically create labeled training data. For example, the tweet *"And all I need is one fuckin sheet stamped! #rage"* will be labeled with emotion *anger* and added into training data after removing the ending emotion hashtag "#rage". In this way, we collected a large number of self-labeled tweets covering seven emotions: *joy*, *sadness*, *anger*, *love*, *fear*, *thankfulness* and *surprise*. We collected about 2 million tweets for training, 250 thousand tweets for testing and nearly 250 thousand tweets for algorithm development, all of which were used in our experiment. We did not apply all the features used in [29], instead, we applied a combination of unigram, bigram and LIWC [3] features. LIWC features refer to the percentages of positive and negative emotion words according to LIWC dictionary. This combination achieved a reasonably good accuracy, very close to the best performance achieved by incorporating more features, according to the feature engineering experiments (Refer to Table III in [29]).

We train seven binary classifiers for seven emotions, such that for each emotion e_i, the corresponding classifier C_{e_i} predicts the probability p_j of a tweet t_j expressing the emotion e_i: $p_j = C_{e_i}(t_j)$. Specifically, we train a binary classifier C_{e_i} by selecting all the tweets of a specific emotion (e.g., *anger*) and randomly selecting the same number of tweets that do not express this emotion (the tweets may express other emotions such as *sadness*, *love*, etc.) For a given tweet t_j, we apply all the seven classifiers. If a classifier C_{e_i} provides the highest probability that t_j expresses the emotion e_i among all the classifiers, and this probability is greater than or equal to the predefined threshold τ, we conclude that the emotion e_i is expressed in t_j; otherwise, t_j is labeled as *other*.

$$\begin{cases} Emotion\ e_i & i = \arg\max_k \{C_{e_k}(t_j)\}\ and\ C_{e_i}(t_j) >= \tau \\ Other & Otherwise \end{cases} \tag{1}$$

[3]`http://www.liwc.net`

	Mehl 2003 et. al. [12]	Subrahmanyam 2006 et. al. [18]	**Twitter**
Subject	52 undergraduates	1,150 chatroom users	14 million Twitter users
Sample	4 days' tape recording	12,258 utterance	51 million tweets
Cursing Frequency	0.5% of all words	3% of all utterances	1.15% of all words 7.73% of all tweets

Table 2. Cursing frequency over different datasets: Cursing on Twitter is much more frequent than that in the other two datasets – 1.15% of all words vs. 0.5% of all words, and 7.73% of all tweets vs. 3% of all utterances

Intuitively, higher the value of τ, higher is the precision of identifying the seven emotions, at the expense of recall. To find a τ that provides high precision and reasonable recall, we tried a series of τ values on the development dataset: starting from 0, with an increment of 0.02, and ending at 1.0. We plot the precision and recall of individual emotion classifiers as well as the combined classifier in Figure 3. As we can see from the figure, with the increasing value of τ, the precision increases, while the recall decreases. Emotion classifiers that are on the upper right perform better than those on the lower left. Since we are interested in only the emotions that we can predict with high precision, we skip detecting emotion *surprise* and *fear*, for which the highest precision is less than 65%. We select $\tau = 0.88$ for later emotion identification, with which the combined classifier achieves good a precision while retaining a reasonable recall among all the values of τ we have tested using the development dataset.

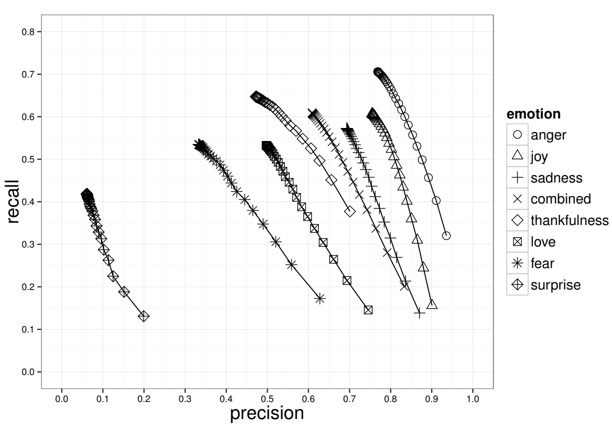

Figure 3. Performance of emotion identification on the development dataset

We then train the classifiers on the training dataset and apply the combined classifiers to the testing dataset. The results are shown in Table 3. Among all the five emotion categories, precision ranges between 56.04% (*thankfulness*) and 84.66% (*anger*), recall ranges between 37.22% (*love*) and 57.01% (*thankfulness*), and F1-score ranges between 45.86% and 68.11%. The combined classification achieves a micro-averaged precision of 76.17%, a micro-averaged recall of 46.07%, and a micro-averaged F1-score of 57.41%. This performance is quite reasonable considering that it is a multi-class classification problem.

Finally, we apply the combined classifier to the 51 million cursing tweets, and obtain the emotion distributions on both cursing and non-cursing tweets, which is shown in Figure

Emotions	Precision (%)	Recall (%)	F1-score (%)
anger	84.66	56.97	68.11
joy	82.77	44.81	58.14
sadness	76.05	39.34	51.86
love	59.72	37.22	45.86
thankfulness	56.04	57.01	56.53
combined	76.17*	46.07*	57.41*

Table 3. Performance of emotion identification on the testing dataset. * micro-averaged metrics. (*Surprise* and *Fear* were dropped because we couldn't detect it with a reasonably high precision on the development dataset)

4. Not surprisingly, cursing is associated with negative emotions: 21.83% and 16.79% of the cursing tweets express sadness and anger emotions, respectively. In contrast, 11.31% and 4.50% of the non-cursing tweets express sadness and anger emotions, respectively. This can be explained by the fact that curse words are usually used for venting out negative emotions, especially anger and sadness. However, we also find that 6.59% of cursing tweets express love. One reason is that curse words can be used to emphasize emotions, including positive ones such as love: e.g., *"**fucking** love you."* Another reason is that certain curse words are used between close friends as a playful interaction, e.g., close female friends call each other *whore*. To better understand how curse words are used to express emotions in tweets, we list some example cursing tweets in Table 4.

In addition, we also examined the frequency of cursing in each type of emotional tweets. As we expected, 23.82% angry tweets and 13.93% sad tweets contain curse words, which are much higher than the rate of curse words in other emotional tweets, such as love (4.16%), thankfulness (3.26%) and joy (2.5%), or the remaining tweets that are not labeled with any of these five emotions (6.39%). This again shows that curse words are often used to express negative emotions.

Cursing vs. Time

Previous study [10] has shown marked difference in emotions (e.g., stress, happiness) expressed between weekdays and weekend, or between morning and night. Similarly, we investigate the relationship between cursing and two types of time periods: times during a day and days of a week. For each tweet, Twitter provides a timestamp based on UTC timezone, indicating when the tweet was posted. However, it makes more sense to use local time when the tweet was posted, so we calculate the corresponding local timestamp for every tweet whose sender has specified timezone in his/her profile. In Figure 5, the lines with triangles and crosses stand for the volumes of overall tweets and cursing tweets, and the line with circles stands for the ratio of cursing tweets to overall

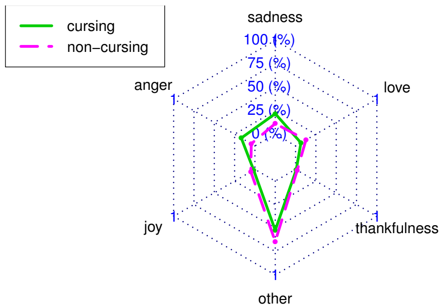

Figure 4. Emotion distributions in both cursing and non-cursing tweets. It shows that curse words are usually used for venting out negative emotions: 21.83% and 16.79% of the cursing tweets express sadness and anger emotions, respectively; in contrast, 11.31% and 4.50% of the non-cursing tweets express sadness and anger emotions, respectively

sadness	"Where da **fuq** is the sun at, this weather is so #depressing" "My life fell apart a long **ass** time ago.. So everythings normal i guess."
anger	"Soo **pissed** off" "People laugh when I say I work at McDonald's. And I say, **bitch** at least I have a job! At least I don't bother my parents asking them for $$$"
love	"Across the ocean, across the sea starting to hate the **fucking** distance between Justin Bieber and me." "@user you little **whore** TAKE ME WITH YOU" "Dear Marilyn Manson, I **fucking** love you and your music. The end."

Table 4. Example tweets in which curse words are used to express different emotions.

tweets. A flat segment of the line with circles suggests the cursing ratio is stable – the increment of cursing tweets keeps pace with that of overall tweets. A rising line segment with circles suggests that the increment of cursing tweets outpaces that of overall tweets. A falling line segment with circles suggests that the increment of cursing tweets is outpaced by that of overall tweets.

We have the following interesting observations from Figure 5. First, the pattern of overall tweet volume fits humans' diurnal activity schedule: it starts rising at 5 am when people get up at the beginning of a day. From then, it keeps rising, and reaches a small peak around lunch time. It keeps rising until it reaches the peak of the day around 9 pm, after which people start preparing to go to sleep. Second, cursing is ever-lasting: the black cursing ratio line with circles always stays above 0, suggesting that people curse all the time throughout the day. Third, the increment of cursing outpaces the increment of overall tweet volume during most of the day time: people curse more and more as they go through the day! In particular, there are two sharp rising slopes: 6 am - 11 am and 3 pm - 1:30 am. We speculate that Twitter users being in good mood during lunch contributes to the flat ratio line segment between 11 am - 2 pm (lunch time). It seems that midnight to 1:30 am is the *high time* for cursing. After that, the volume of cursing tweets decreases faster than the overall tweets.

We now explore the popularity changes of top seven curse words (refer to Figure 1) at different times of a day to gain more insights. We define *relative frequency* for a curse word as its total number of occurrences in any tweet divided by the total number of tweets in a predefined time window. Three representative time windows are selected: 12 am - 2 am, 5 am - 7 am and 12 pm - 2 pm. We observe that the relative frequencies for all top seven curse words keep increasing from 5 am - 7 am to 12 pm - 2 pm and from 12 pm - 2 pm to 12 am - 2 am. On average, from 5 am - 7 am to 12 am - 2 am, the relative frequencies of top seven curse words have increased by 144.05%. In descending order of their relative increase of relative frequencies, top seven curse words rank as follows: *nigga* (262.55%), *ass* (200.70%), *bitch* (199.61%), *shit* (118.86%), *whore* (91.74%), *fuck* (81.93%) and *hell* (52.98%).

To explore how people curse during different days of a week, we plot the ratio of cursing tweets to total tweets each day for four weeks, separately, in Figure 6. The general trend is that users start with relatively high cursing ratios on Mondays, Tuesdays and Wednesdays, then the ratios keep decreasing on the following three days, and reaches the lowest point on Saturdays. Then they start rising up on Sundays. To see the general trend clearly, readers are referred to see the four-week average ratio in the plot. Although we observe this general pattern across four weeks, we are still unclear about the reason. We are interested in the popularity changes of top seven words during different days of a week, similar to those at different times of a day. We select the following two time windows: Monday-Tuesday and Friday-Saturday. On average, from Friday-Saturday to Monday-Tuesday, the relative frequencies of top seven curse words have increased by 10.28%. In descending order of their relative increase of relative frequencies, top seven words rank as follows: *bitch* (14.75%), *shit* (13.10%), *nigga* (11.54%), *whore* (9.64%), *ass* (9.40%), *hell* (7.08%) and *fuck* (6.45%).

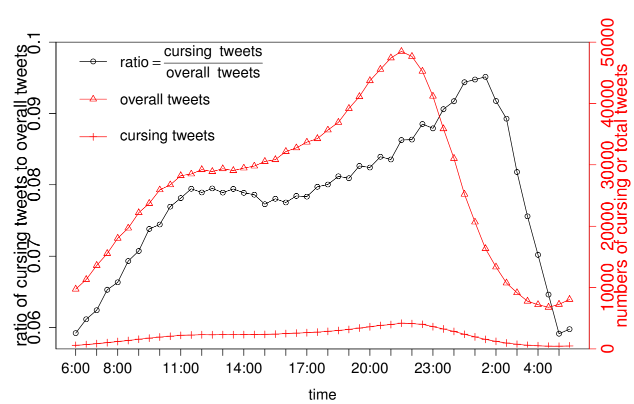

Figure 5. Cursing Volume at Different Times of Day

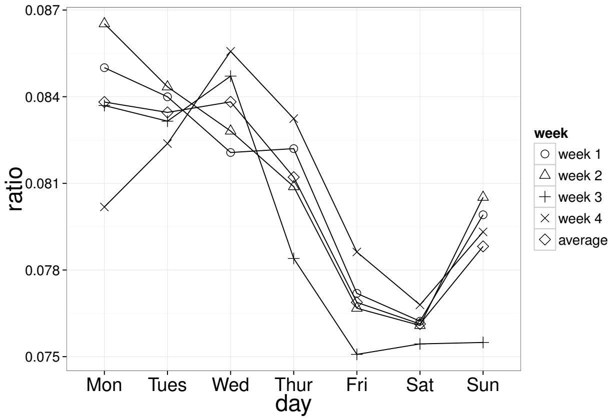

Figure 6. Cursing Volume in Different Days of Week

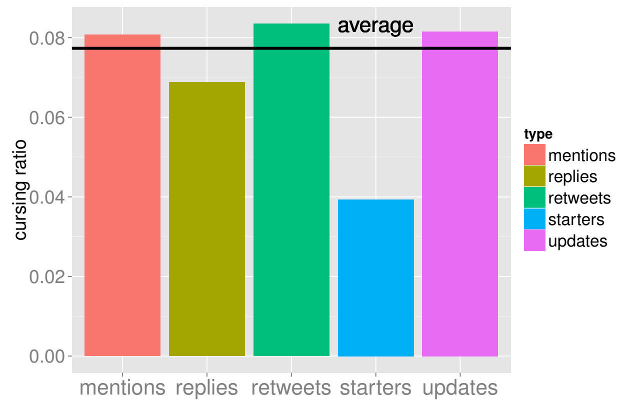

Figure 7. Cursing Ratios in Different Types of Messages

Cursing vs. Message Type

Tweets can be grouped into different message types and we are curious whether users curse differently in different types of tweets. Specifically, *retweet* refers to the tweet that is simply a re-posting of a tweet from another user. If a user receives a tweet from another user, and this user clicks on *reply* button to write a new tweet to reply to this tweet, then this newly posted tweet is called a *reply*. If a user starts sending a tweet to another user, and this tweet is not a reply to any other tweet, we name it a *starter*. If a tweet mentions another user, and it is neither a *reply* nor a *starter*, we call it a *mention*. If a tweet does not belong to any of the above categories, it is an *update*.

We plot the ratio of cursing tweets in each message category in Figure 7, where the black horizontal line stands for the average ratio of cursing tweets to all the tweets. It is interesting to note that although we see quite a bit of cursing messages on Twitter in general, when the messages are sent to other users, the cursing ratios are below average. The ratio of cursing tweets in *starters* is 3.93%, which is only 51.01% of the average cursing ratio. This suggests that users perform self-censorship to some extent when they directly talk to other users. When they post updates about themselves or simply mention other users' names, they do not pay as much attention to the use of curse words. Another interesting observation is that the highest cursing ratio occurs in *retweets*. Sood et. al. [16] found that "*profane comments are more popular or more widely read than non-profane comments*" by receiving thumb ups and downs in Yahoo! Buzz. We are interested in assessing whether use of curse words can help draw other users' attention so as to be retweeted. However, Pearson's product-moment correlation analysis between whether a tweet has curse words and the number of times it is retweeted suggests a negligible correlation: $r = -0.00154, p < 2.2e - 16$. We perform the same analysis on whether a tweet has curse words and whether it is retweeted, and find a stronger but still very weak positive correlation: $r = 0.03366, p < 2.2e - 16$. Similarly, a negligible correlation is also observed between whether a tweet has curse words and whether it has been favorited: $r = 0.005436, p < 2.2e - 16$.

Cursing vs. Location

Location also affects the way people curse in physical world. Cameron [3] found that people curse more at parties than they do at work places. Jay [4] has a similar discovery obtained by investigating cursing frequency at different campus locations: people tend to curse more in relaxed environments (e.g., college dorms vs. Dean's office). Compared with physical world conversations, tweets are posted in digital world: a user can use curse words without being noticed by surrounding people. Does physical locations still affect Twitter users' cursing frequency? Luckily, a Twitter user's location can be inferred via geo-enabled tweet feature. This feature provides the latitude and longitude of the user's location, along with the usual tweet content.

Given a pair of latitude and longitude, Foursquare's venue search API [4] returns a list of nearby venues, as well as the distances. If the distance from the user's location to the nearest venue is less than 50 meters, we assume the user posted the tweet from that venue. For every venue, we retrieve its immediate category and upper category from Foursquare, e.g., a venue is under Asian Restaurant (immediate category), and Asian Restaurant is under Food (upper category). We made very few changes to the Foursquare category hierarchy to reduce ambiguous information, e.g., we deleted *other places* from the hierarchy, because we have no idea what the name suggests. We also removed categories if they have very sparse tweets. Table 5 shows the different categories of venues, the raw number of cursing tweets and the ratio of cursing tweets to all the tweets sent from venues of the same category.

We have the following observations: a) The pattern of more swearing in more relaxed environment still holds, e.g., cursing ratios in a descending order are: Residence (7.08%) > Shop & Service (6.41%) > Nightlife Spot (6.37%) > Entertainment & Recreation (5.71%) > Professional Places (5.64%) > Travel & Transport (5.34%). However, the gaps are much less than those in physical world, partly due to the fact that communications happen in digit world. b) Two exceptions, College Academic Place and High School, have

[4] `https://developer.foursquare.com/docs/venues/search`

Venues	Cursing Tweets (#)	Cursing Tweets (%)
Field (Nature)	380	**4.97**
Travel & Transport	621	5.34
Food	2814	5.35
Professional Places	2020	5.64
Entertainment & Recreation	1305	5.71
Arts	195	5.77
Nightlife Spot	1063	6.37
Shop & Service	3036	6.41
College Academic Place	1155	6.45
Residence	2198	7.08
High School	339	**9.36**

Table 5. Cursing ratios from different places. Field: lakes, Beach, Mountain, etc.; Travel & Transport: Train, Plane, Ferry, etc.; Professional Places: Police Station, City Hall, Office, etc.; College Academic Place: Law School, Engineering Building, Math Building, etc.; Residence: Home, Residential Building, Hotel, etc.

very high cursing rates. This suggests that young high school and college students tend to use more curse words, even in educational places. c) We speculate that users are usually in a good mood while out in the nature, and that is why its cursing ratio is the lowest (4.97%) among all the venues.

Cursing vs. Gender

Another interesting question about cursing is "*who says curse words to whom*". We first explore the gender factor in this section, and then discuss the effect of social rank in the next section. Prior studies have found that gender affects the cursing frequency and the choice of curse words, in addition, people curse more in same-gender contexts than in mixed-gender contexts [12, 11, 8, 15]. We explore whether these hypotheses still hold when people send messages to each other on Twitter. In order to study the gender difference, we first apply an algorithm to recognize the gender of users in our data collection. A person's gender can be revealed by his/her first name: Linda, Lisa, Betty, etc. are usually females' names; John, Paul, William, etc. are male names. US Census Bureau [5] provides 1,219 most popular male names and 4,275 most popular female names. We calculate the "maleness" or "femaleness" by dividing the number of male/female people using this name by the overall size of the population. If a name has a high female percentage and a low male percentage, e.g., Mary: Male 0.009%, Female: 2.629%, then the corresponding person is mostly female. If the female and male percentages of a name are close, e.g., Morgan: Male 1.8%, Female: 2.2%, it suggests that this name is usually used for both genders. If a name is missing from male (female) name list, we take it as female (male). Our algorithm will label a user as female when the female percentage divided by the male percentage is greater than or equal to four; if the male percentage divided by the female percentage is greater than or equal to four, the user will be labeled as male; otherwise, the user will be labeled as unknown.

Overall, this algorithm identified 4,639,204 females and 3,826,701 males in our Twitter user collection. Recall that previously we grouped tweets into five categories: *mention*, *reply*, *retweet*, *starter* and *update*. Here we consider only

[5] `https://www.census.gov/genealogy/www/data/1990surnames/names_files.html`

Sender	Recipient	Cursing Tweets (#)	Cursing Ratio (%)
F	M	3,808	**3.81**
F	F	3,977	3.98
M	F	4,192	4.19
M	M	5,483	**5.48**

Table 6. Cross Gender Cursing Statistics. Statistics of each row are drawn on randomly sampled 100K tweets. Reported *cursing ratio* in each row is the percentage of cursing tweets out of all the tweets within each corresponding group.

Word	F→F	F→M	M→F	M→M	χ^2
fuck	1236	1284	1359	**2069**	308.89***
shit	670	661	831	**1159**	195.61***
nigga	119	171	201	**338**	126.59***
bitch	**475**	273	281	298	83.46***
hell	334	315	349	**532**	79.43***
dick	54	67	80	**137**	47.49***
cunt	22	24	26	**60**	29.70***
fag	20	29	26	**58**	25.83***
pussy	25	25	33	**60**	23.13***
slut	**50**	30	24	17	19.99***
ass	**1091**	1030	970	922	16.07**
bastard	9	14	18	**32**	16.04**
piss	95	91	107	**143**	15.41**
cock	14	15	25	**37**	15.15**
whore	**92**	68	68	50	12.82**

Table 7. The frequency of curse words out of 100K tweets posted or received by males and females. *** $p \leq 0.001$, ** $p \leq 0.01$

reply and *starter*, since they represent targeted messages between Twitter users with explicit message sender (*who*) and recipient (*to whom*) specified. These messages are further divided into four groups based on gender – *female to female*, *male to female*, *female to male* and *male to male*. To make results comparable, we randomly sampled 100K tweets from each of these four groups and statistics are shown in Table 6.

Comparing the same-gender contexts (*F to F* and *M to M*) with the mix-gender contexts (*F to M* and *M to F*) in Table 6, we observe that people are more likely to use curse words within the same-gender context, and this tendency is more obvious when the message senders are males (5.48% vs. 4.19%). This is consistent with the findings in prior studies [8, 15] on offline communications. Moreover, Male-to-Male communication has the highest cursing ratio: 5.48%, while Female-to-Male has the lowest cursing ratio: 3.81%.

Regarding the preference of curse words, out of randomly sampled 100K tweets for each of the four groups (see Table 7), we also find clear difference between females and males. There are a set of words that are used significantly more often by males than by females, for example, *fuck*, *shit*, and *nigga*. Some other words are significantly overused by females, such as *bitch* and *slut*. It is also interesting to observe that such differences are more apparent between two same-gender contexts – *F to F* vs. *M to M*. This suggests that the genders of both "*who*" and "*whom*" matter in the choice of curse words.

Cursing vs. Social Rank

We now look into the relationship between social rank and cursing behavior. Within a society, it is expected that higher the social rank of a person, the less cursing the person performs [8]. We use the number of followers on Twitter as

User Group	Sender				Recipient			
	Cursing Tweet (#)	Cursing Ratio(%)	$\mu_{followers}$	$\sigma_{followers}$	Cursing Tweet (#)	Cursing Ratio(%)	$\mu_{followers}$	$\sigma_{followers}$
top 1%	146,035	5.98	67,810	408,228.8	49,069	3.91	155,000	650,825.3
1% - 10%	847,467	**8.78**	1,923.00	1,481.60	101,983	6.11	3,764.00	3,744.18
10% - 40%	1,744,258	**8.75**	400.10	148.20	289,035	**7.96**	565.70	219.82
40% - 90%	1,116,645	6.62	101.60	60.58	258,984	6.26	172.20	71.85
90% - 100%	77,523	4.00	2.30	2.91	28,348	4.56	29.91	15.77

Table 8. Cursing Ratio vs. Social Ranking (followers) for both Senders and Recipients. μ population mean, σ standard deviation. For senders and recipients, we show statistics regarding their posted and received tweets, respectively

Word	top 1%	1-10%	10-40%	40-90%	90-100%	$\chi^2 ***$
fuck	2621	3306	**3399**	2814	1265	1093.81
ass	744	**1624**	1607	1027	675	738.11
nigga	563	**1354**	1131	564	696	603.74
shit	986	1588	**1614**	1221	668	534.82
bitch	779	**1224**	1095	763	596	301.16
cock	30	24	19	22	**129**	199.26
blowjob	24	16	15	16	**89**	128.56
dick	178	**203**	169	130	64	78.92
piss	98	119	**159**	148	54	60.98
whore	167	**205**	183	136	95	46.85
pussy	151	**160**	117	76	101	40.18
hell	286	358	**375**	357	253	34.73
slut	**71**	41	50	54	25	23.79

Table 9. The frequency of curse words out of 100K tweets based on the social rank (follower counts) of senders. χ^2 results are based on the comparison of frequencies of each word across different sender groups. *** p≤0.001 for all the values in this column

Word	top 1%	1-10%	10-40%	40-90%	90-100%	χ^2
ass	618	1521	**2284**	1590	1034	1119.08***
nigga	243	674	**892**	471	266	599.30***
shit	628	1089	**1428**	1090	774	387.95***
fuck	1556	1875	**2330**	2023	1507	250.08***
bitch	350	533	**643**	458	346	136.78***
hell	244	431	**522**	434	370	105.54***
pussy	62	**88**	71	52	39	22.20***
whore	92	128	**145**	102	92	20.08***
slut	**56**	27	32	35	24	18.24**
dick	124	133	**134**	110	84	14.80**
piss	65	80	96	**107**	93	11.82*

Table 10. The frequency of curse words out of 100K tweets based on the social rank (follower counts) of recipients. χ^2 results are based on the comparison of frequencies of each word across different recipient groups. *** p≤0.001, ** p≤0.01 * p≤0.05

an approximation to the social rank in digital world. We sort both senders and recipients based on the descending order of their number of followers, and then divide them into five groups: top 1% (who have the highest numbers of followers), then followed by 1% - 10%, 10% - 40%, 40% - 90%, 90% - 100%. In Table 8, we show the raw numbers of posted/received cursing tweets, the ratio of posted/received cursing tweets out of overall tweets, the mean and standard deviation of followers that the group users have, within each sender/recipient group.

The top 1% senders do curse, but it is less than what we expected. We also observe bell-shaped distributions in cursing ratios for both senders and recipients: the middle sender groups (1%-10% and 10%-40%) curse the most, while the middle recipient group (10%-40%) receive tweets with the highest cursing ratio. Senders from the bottom group, who may have recently joined Twitter, and have very few followers (mean: 2.3), curse the least among all sender groups. Turning to recipients, the cursing ratio among tweets received by the top 1% group, is the lowest across all recipient groups: these popular users receive a lot of friendly messages from their fans, e.g., "*@Harry_Styles follow me babe<3*", *"@NiallOfficial I can't sleep :("*

Besides the cursing ratios in tweets that are posted/received by different user groups, we are also interested in the curse word choices across all groups. To make the results comparable, we randomly sampled 100K posted tweets from each sender group and counted the corresponding frequencies of curse words in Table 9. We did the same to all the tweets received by different recipient groups in Table 10. We observed that the same word can be used at different rates across groups: 10-40% sender group used *fuck* 3,399 times out of 100K posted tweets, while 90-100% sender group used it only 1,265 times; 10-40% recipient group received *ass* 2,284 times out of 100K received tweets, while top 1% recipient group only received it 618 times. We find that, for the same word, its post/received volumes usually achieve highest frequencies in 1-10% and 10-40% for sender groups, and 10-40% for recipient groups with a few exceptions. The reason why top 1% sender group used *slut* word the most is because there are a few popular Twitter accounts that posted funny tweets about the word *slut*. We also found some porn accounts in 90-100% sender group that aggressively posted porn links, which explains the peaks for the words *cock* and *blowjob* in this group. The reason why top 1% recipient group received more tweets containing *slut* is because some fans like to call celebrities *slut* regardless of their gender for fun, e.g., *"@taylorswift13 slut"*, *"@Harry_Styles slut drop on my follow button :))))))))"*

LIMITATIONS

This study is limited in several ways. Firstly, our exploration is based on a random sampling of tweets posted by Twitter users, and our results may be biased towards these users, who may not statistically represent users in other social media websites or mirror overall population in the real world. Thus our findings may turn out to be different on other datasets or social media platforms. Though the findings may not be generalized beyond Twitter, the analysis framework can be applied to cross-platform studies which we would like to pursue in our future work. Secondly, with such a large amount of data, it is impossible to manually label all the tweets/users. Because of this, we employed automatic approaches to labeling, profanity, emotion, gender, etc. Though these techniques have been used in many studies, they are not perfect. It is important to note that we always choose precision over recall when designing these techniques. Further improving any of these approaches would be an interesting topic for future research. Moreover, in a few tasks, we relied on self-reported data, such as users' names and geo-locations. Using self-reported data may lead to bias toward users who opt to share such information. Further, we do not segment Twitter user accounts into different types, e.g., celebrity, media, organization or regular personal accounts. Users from these different categories may curse in varied manner, and it would be interesting to examine the differences. Other topics as for extension of this study are the comparative study of native speakers and non-native speakers in using curse words, and the explorations of cursing behavior in languages other than English. Finally, this work is mostly descriptive, which provides insights on the *what* aspect of the cursing phenomenon on Twitter. In order to achieve deeper understanding on *why*, e.g., why people choose particular curse words they use, why people's cursing behaviors depend on certain contextual variables, etc., user surveys and qualitative analysis will be needed.

CONCLUSION

In this paper, we investigated the use of curse words in the context of Twitter based on the analysis of randomly collected 51 million tweets and about 14 million users. In particular, we explored four questions that have been identified as important by the prior swearing studies in the areas of psychology, sociology, and linguistics.

Regarding the question of *ubiquity* of cursing on Twitter, we examined the frequency of cursing and people's preference in the use of specific curse words. We found that the curse words occurred at the rate of 1.15% on Twitter, and 7.73% of all the tweets in our dataset contained curse words. We also found that seven most frequently used curse words accounted for more than 90% of all the cursing occurrences. The second question we studied is the *utility* of cursing, especially the use of cursing to express emotions. We built a classifier which identified five different emotions from tweets – *anger*, *joy*, *sadness*, *love*, and *thankfulness*. Based on the classification results, we found that cursing on Twitter was most closely associated with two negative emotions: *sadness* and *anger*. However, curse words could also be used to emphasize positive emotions such as *joy* or *love*.

Prior studies suggest that cursing is sensitive to various contextual variables. We focused on examining three contextual variables regarding *when*, *where* and *how* the cursing occurs. We found that the pattern of overall tweet volume matches peoples diurnal activity schedule, and people curse more and more after they getup in the morning till sleep hours of the night. Our study of the relation between cursing and message types suggests that users perform self-censorship when they talk directly to other users. We find that users do curse more in relaxed environments, but the differences across different environments are very small, partly due to the fact that Twitter messages are posted in virtual digital world.

The last question we tried to investigate is about *who says curse words to whom*. We examined the gender and social rank factors and how they might affect people's cursing behaviors on Twitter. Our results support the findings from prior studies that gender and social rank relate to people's propensity to curse and the choice of curse words. Specifically, men curse more than women, men overuse some curse words different from what women use and vice versa, and both men and women are more likely to curse in the same-gender contexts. Turning to social rank, high rank users do curse less than most low rank users; the ratios of using/receiving curse words achieve highest numbers in 1%-10%, 10%-40% sender groups, and 10%-40% recipient group.

ACKNOWLEDGMENTS

This research was supported by US National Science Foundation grant IIS-1111182: SoCS: Social Media Enhanced Organizational Sensemaking in Emergency Response. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

1. Almuhimedi, H., Wilson, S., Liu, B., Sadeh, N., and Acquisti, A. Tweets are forever: a large-scale quantitative analysis of deleted tweets. In *Proceedings of the 2013 conference on Computer supported cooperative work*, ACM (2013), 897–908.
2. Bak, J. Y., Kim, S., and Oh, A. Self-disclosure and relationship strength in twitter conversations. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2*, Association for Computational Linguistics (2012), 60–64.
3. Cameron, P. Frequency and kinds of words in various social settings, or what the hell's going on? *The Pacific Sociological Review 12*, 2 (1969), 101–104.
4. Jay, T. *Cursing in America: A Psycholinguistic Study of Dirty Language in the Courts, in the Movies, in the Schoolyards, and on the Streets*. John Benjamins Publishing Co, 1992.
5. Jay, T. *Why we curse: A neuro-psycho-social theory of speech*. John Benjamins Publishing, 2000.
6. Jay, T. The utility and ubiquity of taboo words. *Perspectives on Psychological Science 4*, 2 (2009), 153–161.
7. Jay, T., et al. Do offensive words harm people? *Psych. Pub. Pol. and L. 15* (2009), 81–315.
8. Jay, T., and Janschewitz, K. The pragmatics of swearing. *Journal of Politeness Research. Language, Behaviour, Culture 4*, 2 (2008), 267–288.
9. Jay, T. B. Sex roles and dirty word usage: a review of the literature and a reply to haas. *Psychological Bulletin 88*, 3 (1980), 614–621.
10. Kamvar, S. D., and Harris, J. We feel fine and searching the emotional web. In *Proceedings of the fourth ACM*

international conference on Web search and data mining, ACM (2011), 117–126.

11. McEnery, T. *Swearing in English: Bad language, purity and power from 1586 to the present*, vol. 1. Psychology Press, 2006.
12. Mehl, M. R., Pennebaker, J. W., et al. The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of personality and social psychology 84*, 4 (2003), 857–870.
13. Morris, M. R., Counts, S., Roseway, A., Hoff, A., and Schwarz, J. Tweeting is believing?: understanding microblog credibility perceptions. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM (2012), 441–450.
14. Nasution, R. F., and Rosa, R. N. Swearwords found in chat room yahoo messenger (page: 84-93). *E-Journal English Language and Literature 1*, 1 (2012).
15. Pilotti, M., Almand, J., Mahamane, S., and Martinez, M. Taboo words in expressive language: Do sex and primary language matter? *American International Journal of Contemporary Research 2*, 2 (2012), 17–26.
16. Sood, S., Antin, J., and Churchill, E. Profanity use in online communities. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, ACM (2012), 1481–1490.
17. Stephens, R., Atkins, J., and Kingston, A. Swearing as a response to pain. *Neuroreport 20*, 12 (2009), 1056–1060.
18. Subrahmanyam, K., Smahel, D., Greenfield, P., et al. Connecting developmental constructions to the internet: Identity presentation and sexual exploration in online teen chat rooms. *Developmental psychology 42*, 3 (2006), 395.
19. Thelwall, M. Fk yea i swear: Cursing and gender in a corpus of myspace pages. *Corpora 3*, 1 (2008), 83–107.
20. Banbuilder - PHP curse word function and word list for application developers, moderators, etc., March 2013. `http://banbuilder.com/`.
21. BannedWordList.com - a resource for web administrators, March 2013. `http://www.bannedwordlist.com/`.
22. Habitat for Profanity, July 2010. `http://www.parentstv.org/PTC/publications/reports/2010ProfanityStudy/study.pdf`.
23. List of Bad Words, March 2013. `http://www.noswearing.com/dictionary/`.
24. How American teens navigate the new world of "digital citizenship", November 2011. `http://pewinternet.org/Reports/2011/Teens-and-social-media/Summary/Findings.aspx`.
25. Pew Internet: Social Networking (full detail), Febrary 2013. `http://pewinternet.org/Commentary/2012/March/Pew-Internet-Social-Networking-full-detail.aspx`.
26. Profanity - Wikipedia, the free encyclopedia, March 2013. `http://en.wikipedia.org/wiki/Profanity`.
27. Watch your mouth! Americans see profanity getting worse, poll finds, March 2006. `http://www.nbcnews.com/id/12063093/from/RS.4/?GT1=7938#.UaKaG2cSpZo`.
28. Bad words list, March 2013. `http://urbanoalvarez.es/blog/2008/04/04/bad-words-list/`.
29. Wang, W., Chen, L., Thirunarayan, K., and Sheth, A. P. Harnessing twitter" big data" for automatic emotion identification. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, IEEE (2012), 587–592.
30. Xiang, G., Fan, B., Wang, L., Hong, J., and Rose, C. Detecting offensive tweets via topical feature discovery over a large scale twitter corpus. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, ACM (2012), 1980–1984.
31. Xu, J.-M., Zhu, X., and Bellmore, A. Fast learning for sentiment analysis on bullying. In *Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining*, ACM (2012), 10.